

Teljes szövegű publikációk cseréje repozitóriumok között - a SWORD protokoll tesztelése az MTA Könyvtárában és a Debreceni Egyetem Könyvtárában

Holl András

MTA Könyvtára, MTA K-TM Csillagászati Kutatóintézete

Karácsony Gyöngyi, Harangi Gabriella

Debreceni Egyetem Egyetemi és Nemzeti Könyvtár

Cserbák Márton

KFKI Rendszerintegrációs Zrt

Vetített anyag: <http://www.konkoly.hu/staff/holl/sword/swordsli.pdf>

Tartalmi kivonat

A Budapesti Nyílt Hozzáférés Kezdeményezés valamint a Berlini Deklaráció nyomán, a Nyílt Archívumok Kezdeményezés keretében kidolgozott protokoll és szoftverek segítségével tudományos publikációkat tároló és hozzáférhetővé tevő repozitóriumok keletkeznek. Ezek lehetnek intézményiek, tudományterület vagy kutatásfinanszírozó ügynökség szerint szerveződők. Egyre több intézmény és szervezet kötelezi dolgozóit vagy az általa támogatott kutatókat cikkeik repozitóriumban való elhelyezésére. Előfordulhat, hogy egy kutatónak vagy kutatócsoportnak több elhelyezési kötelezettségnek is eleget kell tennie. Ilyen esetben szerencsés, ha a szerzőknek elegendő a cikküket egy repozitóriumban elhelyezniük, és a repozitórium a teljes szövegű cikket automatikusan továbbítani tudja további illetékes repozitórium vagy repozitóriumok felé. Erre a célra szolgál a SWORD protokoll.

Cikkünkben az MTA Könyvtárában és a Debreceni Egyetem Egyetemi és Nemzeti Könyvtárában tesztelt, SWORD alapú adatsere megoldást mutatjuk be.

*

Közismert a tudományos eredmények publikációs problémája: a közpénzből (adóból) támogatott kutatások eredményeit kiadóvállalatok jelentetik meg folyóiratokban, a cikkek szerzői és bírálói honoráriumot nem kapnak, a (szerzőket foglalkoztató) intézmények pedig igen magas áron vásárolhatják meg a folyóiratokat. A kutatásértékelési rendszerben az impakt faktor, mint a folyóiratokat jellemző mérőszám¹ vagy az SCI idézetek használata, valamint a könyvtárak azon törekvése, hogy minden, a gyűjtőkörükbe tartozó folyóiratot beszerezzenek, megnehezíti az említett helyzet megváltoztatását.

1 Néha - helytelenül - a kutatók teljesítményének jellemzésére is használják.

A megoldás lehetőségét az Open Access - a nyílt hozzáférés - elv alkalmazása jelenti. A tudományos folyóiratcikketeket vagy eleve szabad hozzáférést biztosító folyóiratokban kell közölni (ez az "arany út"), vagy szabadon hozzáférhető repozitóriumokban is el kell helyezni ("zöld út"). Repozitóriumok és OA folyóiratok évek óta léteznek, a szabadon elérhető cikkek részaránya mégsem haladja meg összességében a 20%-ot (Björk, Roos & Lauri, 2009).

A cikkek repozitóriumban való elhelyezésének alacsony szintjét okozhatja a szerzők nehézkessége és tudatlansága. A helyzetet javíthat, ha a munkáltatójuk, vagy a kutatást finanszírozó szervezet megköveteli a nyílt hozzáférést. A cikkek repozitóriumba való feltöltése nem bonyolult, ám ennek további egyszerűsítése is jótékony hatással lehet az elhelyezések arányára nézve. Különösen abban az esetben lényeges ez, ha a szerző - kötelezettségből vagy valamiféle előny érdekében - több repozitóriumba is elhelyezi ugyanazt a cikket.

Hagyományosan a cikket maga a szerző tölti fel a repozitóriumba egy webes felületen keresztül. A SWORD olyan alkalmazásprogramozási felület (application programming interface, API), mely lehetővé teszi a repozitóriumok cikkekkel való feltöltését programok segítségével. A SWORD-dal megvalósítható, hogy teljes szövegű publikáció, a szükséges metaadatokkal együtt egyetlen egyszerű művelettel - mondjuk egy ikon elhúzásával és egy másik ikonra ejtésével - feltölthető legyen a repozitóriumba, avagy egyetlen űrlap kitöltésével több repozitóriumba is bekerülhessen. Vannak érvek a mellett is, hogy ugyanaz az anyag több repozitóriumba is bekerüljön: redundancia, biztonság, az egyes rendszerek eltérő funkcionalitásának kihasználása. A SWORD segíthet abban, hogy repozitóriumok valamiféle gyarapítási koncepció szerint egymásnak átadjanak teljes szövegű anyagokat.

A SWORD (feloldása: Simple Web-service Offering Repository Deposit) egy, a brit Joint Information Systems Committee által támogatott projekt, melynek első szakasza 2007-ben zajlott. A program javaslatot dolgozott ki egy repozitórium-API megvalósítására. Az Atom Publishing Protocol² profiljaként alakították ki a SWORD-ot: HTTP alapon lehet információkat kérni a repozitóriumról (service document), majd fájlokat feltölteni. A projekt során implementálták a SWORD-öt négy elterjedten használt repozitóriumba (DSpace, EPrints, IntraLibrary és Fedora), demonstrációs SWORD klienseket készítettek, és mindezek forráskódját hozzáférhetővé tették (Allinson, Francois és Lewis, 2008; Currier, 2009).

Mára a SWORD-öt implementálta az első repozitórium, az arXiv³, és beépítették az EPrints 3.1-es és a DSpace 1.5.2-es változatába. Készítettek MS Word és FaceBook klienseket. Megvizsgálták az OAI-ORE⁴ alkalmazásának lehetőségét a SWORD-dal együtt, összetett objektumok repozitóriumba helyezése kapcsán.⁵

2 <http://atompub.org/rfc4287.html>

3 <http://www.swordapp.org/sword/case-studies/sword-case-study-arxiv>

4 <http://www.openarchives.org/ore/>

5 <http://www.openarchives.org/ore/RepoCamp2008/#ORESWORD>

Az MTA Könyvtára és a Debreceni Egyetem Egyetemi és Nemzeti Könyvtár kísérletet végzett a SWORD protokoll felhasználhatóságára vonatkozóan. Az MTA Könyvtárának repozitóriuma (REAL - Repository of the Academy's Library) várhatóan 2009 közepén nyílik meg. Az első gyűjteménye az OTKA támogatásával született tudományos publikációkat és OTKA kutatási jelentéseket tartalmaz majd, és EPrints szoftvert használ. A Debreceni Egyetem Elektronikus Archívuma (DEA) 2006-ban indult el, rektori mandátum alapján tárolja az egyetem dolgozóinak publikációit, jelenleg nagyjából 8500 tételt tartalmaz, és DSpace szoftverrel üzemel.

A kísérlet során a SWORD projekt keretében készült EPrints és DSpace implementációt, és SWORD klienst alkalmaztuk⁶, ez utóbbit mind GUI, mind parancssoros módban. Össze kellett egyeztetni a metaadat-készleteket: a REAL például tárolja az OTKA projekt-azonosítószámokat, a DEA nem. Ugyancsak fontosnak találtuk az egyedi azonosítók (pl. DOI) átadását: ha ugyanaz a tétel több repozitóriumba is bekerül, a duplumszűrést támogatandó egyedi azonosítót is alkalmazni kell. A SWORD a metaadatokat és a dokumentumot valamiféle csomagban fogadja be. Ha a metaadatok és a cikk külön állományokban vannak, ZIP formátumban csomagolhatóak össze. A lehetséges feltöltési formátumok a következők: METS (Guenther, McCallum, 2003) metaadatok és különálló cikk; IMS Content Packaging⁷; EPrints XML (ami base64 kódolással tartalmazza a cikket is). A kísérleteink során kiderült, hogy a különböző SWORD implementációk (EPrints, DSpace) ezek közül a formátumok közül nem mindent képesek kezelni. Az IMS formátummal bonyolultsága, ilyen formátumot kezelni képes szoftverek hiánya miatt nem foglalkoztunk. A DSpace a METS-es formátumot volt képes alkalmazni, az EPrints pedig saját XML export formátumát. A két repozitórium közötti adatcsere megvalósítása úgy lehetséges, hogy az egyik irányban METS formátumú metaadatokat kell létrehozni, és a dokumentummal együtt ZIP-állományba csomagolni, a másik irányban EPrints XML állományokat kell szkriptek segítségével létrehozni. Tapasztalataink szerint a SWORD hibaüzenetei gyakran semmitmondóak.

Véleményünk szerint a SWORD még nem teljesen kiforrott technológia, de már jelenlegi állapotában alkalmazható. Felmerült, hogy az OTKA Iroda a hozzá benyújtott kutatási jelentéseket a SWORD alkalmazásával tölthetné fel a REAL-ba. Tervezzük a kísérletek folytatását a REAL és a DEA között, megvizsgálva annak lehetőségét, hogy a DEA-ba cikket feltöltő felhasználó az OTKA pályázati azonosító megadásával és egyetlen gomb megnyomásával a REAL-ba is elhelyezhesse a feltöltött tételt. Kíváncsok lennének a SWORD protokoll implementálása a Magyar Tudományos Akadémia Köztisztviselői Publikációs Adattárában (MTA KPA) vonatkozásában is.

Irodalom

- B-C. Björk, A. Roos, M. Lauri: Scientific journal publishing: yearly volume and open access availability, 2009, Information Research, 14, No. 1.,

⁶ <http://www.swordapp.org/sword/demonstrators>

⁷ <http://www.imsproject.org/content/packaging/>

<http://informationr.net/ir/14-1/paper391.html>

- J. Allinson, S. Francois, S. Lewis: SWORD: Simple Web-service Offering Repository Deposit, 2008, Ariadne, 54, <http://www.ariadne.ac.uk/issue54/allinson-et-al/>
- S. Currier: Cutting Through the Red Tape to Populate Learning Materials Repositories, 2009, JISC e-Learning Forum, <http://www.elearning.ac.uk/features/sword>
- Guenther, R., McCallum, S.: New metadata standards for digital resources: MODS and METS, 2003, Bulletin of the American Society for Information Science & Technology, <http://www.asis.org/Bulletin/Dec-02/guenthermccallum.html>